August 15, 2024

Via ECF

Hon. Ona T. Wang
United States District Court
Southern District of New York

Re: T*he New York Times Company v. Microsoft Corporation et al.*,
        Case No. 23-cv-11195-SHS-OTW | Dispute Re: OpenAI's Custodians

Dear Magistrate Judge Wang:

        The Court should reject The Times' request because it is premature and seeks relief to which The Times is not entitled.

I.   The Times' motion is premature

    Throughout discovery in this case, The Times has repeatedly burdened the Court with premature discovery motions without completing the required meet and confer process. *See, e.g.*, Dkt. 147; *see also* Dkt. 136. This motion is yet another example. The parties' discussions regarding custodians have been ongoing, and they have not yet reached an impasse. During these discussions, OpenAI initially identified 12 custodians. Later, on July 31, OpenAI offered to add *four* more custodians. The Times never responded to that offer nor made any attempt to meet and confer further before filing this request, without warning, almost two weeks later.

    Had the Times continued to confer with OpenAI—as it was required to do—it would have learned that OpenAI was willing to designate another *ten* custodians, bringing the total to *26*. Notably, those additional custodians include the individuals that feature most prominently in The Times' brief, including Sam Altman, OpenAI's CEO; Greg Brockman, OpenAI's President; Mira Murati, OpenAI's CTO; James Dyett, OpenAI's Head of Platform Sales; and Alex Passos and Johannes Heidecke, members of OpenAI's technical staff. The additions also include another custodian (Peter Welinder) that The Times had previously requested during the parties' meet and confers. As this progress demonstrates, the parties can, and should, continue to meet and confer, as the Court's requirements contemplate.

II.   The Times has not met its burden to show the additional custodians it requests are likely to possess relevant, non-duplicative information

    Although The Times asks the Court to compel OpenAI to add custodians, its letter brief never once mentions the standard for compelling such a search—because The Times cannot meet it. "[T]he responding party is entitled to select the custodians most likely to possess responsive information and to search the files of those individuals. Unless that choice is manifestly unreasonable or the requesting party demonstrates that the resulting production is deficient, the court should play no role in . . . designating custodians." *Mortg. Resol. Servicing, LLC v. JPMorgan Chase Bank, N.A.*, No. 15 Civ. 0293 (LTS) (JCF), 2017 WL 2305398, at *2 (S.D.N.Y. May 18, 2017) (cleaned up). Thus, The Times bears the burden to show that the custodians it requests "would provide unique, relevant and noncumulative evidence." *Coventry Cap. US LLC v. EEA Life Settlements, Inc.*, No. 17 Civ. 7417 (VM), 2021 WL 961750, at *2 (S.D.N.Y. Mar. 15, 2021). The Times cannot clear that bar. It asserts that each of its proposed

2767109

custodians is likely to have "relevant" information. But a requesting party cannot "justify an order compelling Defendants to search the ESI of . . . additional custodians simply by showing that each one had some relation to the events in question." *Enslin v. Coca-Cola Co.*, No. 2:14-CV-06476, 2016 WL 7042206, at *2 (E.D. Pa. June 8, 2016).

Indeed, The Times' brief not only fails to show the additional custodians "would provide *unique*, relevant and *noncumulative* evidence," *Coventry*, 2021 WL 961750, at *2 (emphases added), it affirmatively demonstrates that they are, in fact, cumulative. For example, The Times complains that "OpenAI's custodian list has yielded productions slanted toward [AI model] training." Dkt. 204 at 1. Yet it later asserts, without a hint of irony, that OpenAI should be compelled to designate three new custodians because they know about "*Training Datasets*" and a fourth because she "worked on a project to 'modify [a] *training* objective.'" *Id.* at 3 (emphasis added). The Times has also asked OpenAI to add Ms. Murati as a custodian (which OpenAI has agreed to do) while simultaneously requesting that OpenAI designate her former Chief of Staff (Bianca Martin). *Id.* at 2-3. It then doubles-down by justifying its request for yet another custodian (William Fedus) based on his ████████████████████████ *Id.* at 3. But in neither case does The Times explain why the information it seeks couldn't be gathered from Ms. Murati herself, which makes these additional custodians duplicative. *Id.* Similar deficiencies plague each of the custodians The Times has proposed, as explained below.

OpenAI Leaders: The Times identifies several OpenAI executives and implies they would be appropriate custodians by dint of their "Leadership Roles." *Id.* at 1. But "speculation that [an individual's] position as a senior executive might increase the relevance of his files is not a basis for designating him as a custodian." *Mortg. Resol. Servicing, LLC*, 2017 WL 2305398, at *3. The Times asserts that Ilya Sutskeyer has a "unique perspective on Defendants' models and products," but never attempts to explain what makes that perspective unique. Dkt. 204 at 2. As discussed further below, OpenAI is designating several custodians with knowledge of OpenAI's products and models. Those people include OpenAI's CEO Mr. Altman; its President, Mr. Brockman; and its CTO, Ms. Murati. Beyond that, OpenAI is designating Nick Turley, who leads the ChatGPT product team, and Peter Welinder, OpenAI's VP of Product.

The Times identifies Brad Lightcap as relevant because it asserts he was ████████ ████████████████████████████. But OpenAI is designating James Dyett, the person The Times has identified as taking ████████████████████████ *Id.* at 3. So Mr. Lightcap is cumulative, at best. Rather than justify its demand for duplicative discovery from Mr. Lightcap, The Times' motion mentions him almost in passing, and cites two irrelevant items: (1) ████████████████████████████████ and (2) a video interview that *never* mentions negotiations with The Times. Neither supports granting The Times' request.

Alleged Intentional Infringement and "Irresponsible Conduct": The people The Times has identified for these topics are irrelevant. The Times argues Alex Gray and Ryan Lowe have knowledge ██ ██.

Beyond that, The Times argues that Mr. Gray—who is not a lawyer—has relevant knowledge because he once mused that ████████████████████████████ Dkt. 204 at 3. But even if Mr. Gray's opinion ████████████████████████████ ██.

The Times also identifies Daniel Kokotajlo and William Saunders as relevant to purported

"Irresponsible Conduct," but fails to explain why they would have relevant documents, let alone unique, non-duplicative documents. None of their comments cited by The Times concern model training. Although The Times asserts these comments concern "fair use," it never explains why. Regardless, OpenAI has already designated at least three custodians that have worked on the features that help filter and exclude materials that may be copyrighted. Those custodians include Lilian Weng, OpenAI's Head of Safety Systems, the team responsible for ensuring the safety, robustness, and reliability of OpenAI's models.[1] The other two custodians (Vinnie Monaco and Ian Sohl) are engineers that also worked on OpenAI's filtering and exclusion features. Thus, even if the additional custodians demanded by The Times had relevant documents, those documents would almost certainly be duplicative.

Hallucinations and Regurgitation: The Times also seeks documents from Ms. Martin and Mr. Fedus, whom The Times contends are knowledgeable about ███████████████████ ███████████████████████. But The Times' own exhibits demonstrate that OpenAI's existing custodians (e.g., Ms. Murati, Nick Ryder, Barrett Zoph, and John Schulman) are knowledgeable about those same efforts. *See* Exs. 18, 19, 20. Beyond that, OpenAI is designating at least three other custodians—Mr. Heidecke (whom The Times requested on this topic), Ms. Weng, and Dr. Monaco—that have worked on features ████████████ ████████████████████.[2] Thus, there is no reason to believe that these other individuals would have relevant, non-duplicative documents.

Training Datasets: Finally, as discussed above, The Times seeks more custodians on model training, despite complaining they are already awash in such documents. Dkt. 204 at 1. OpenAI has already designated at least six engineers (Nick Ryder, Alex Paino, Ian Sohl, Chris Koch, Barret Zoph and John Schulman) that specifically work on training the relevant OpenAI models. These engineers are all members of OpenAI's Training teams, and include the heads of the Pre-Training Team (Dr. Ryder) and Post-Training Team (Mr. Zoph), and thus are all specifically and most knowledgeable about training datasets. In light of those custodians, the three people The Times identifies are unlikely to offer any new, relevant information. For two of those individuals (Wojciech Zaremba and James Betker), The Times cannot describe why either would have unique knowledge about the AI models relevant to this case. Instead, they appear either to know about models in general, or models that *aren't* relevant. The third person The Times identifies, Chris Berner, is responsible for OpenAI's "supercomputing clusters," *id.* at 3, which are irrelevant. He also conducts "research in highly efficient model training," but The Times cannot point to any link between that research and its claims. *Id.*

* * *

The 26 custodians OpenAI has already offered are more than sufficient. Indeed, it is more than *twice* the eleven custodians The Times has offered to date. Given this disparity, the Court should not allow The Times to fish for more custodians based on its speculation that those custodians might have more relevant documents. The motion should be denied.

---

[1] https://openai.com/safety-systems/
[2] The Times asserts Ms. Martin ██
██

Respectfully submitted,

/s/ Michelle Ybarra
Michelle Ybarra
Keker, Van Nest & Peters LLP

/s/ Vera Ranieri
Vera Ranieri
Morrison & Foerster LLP

/s/ Elana Nightingale Dawson
Elana Nightingale Dawson
Latham & Watkins LLP

Cc: All Counsel of Record (via ECF)

*The parties use electronic signatures with consent in accordance with Rule 8.5(b) of the Court's ECF Rules.

2767109